

SEQUENCE LISTING

(1) GENERAL INFORMATION

- (i) APPLICANT: Goli, Surya K.
Hillman, Jennifer L.
- (ii) TITLE OF THE INVENTION: A NOVEL GLUTATHIONE S-TRANSFERASE
- (iii) NUMBER OF SEQUENCES: 5
- (iv) CORRESPONDENCE ADDRESS:
 - (A) ADDRESSEE: INCYTE GENOMICS, INC.
 - (B) STREET: 3160 Porter Drive
 - (C) CITY: Palo Alto
 - (D) STATE: CA
 - (E) COUNTRY: US
 - (F) ZIP: 94304
- (v) COMPUTER READABLE FORM:
 - (A) MEDIUM TYPE: Diskette
 - (B) COMPUTER: IBM Compatible
 - (C) OPERATING SYSTEM: DOS
 - (D) SOFTWARE: FastSEQ
- (vi) CURRENT APPLICATION DATA:
 - (A) APPLICATION NUMBER: 09/309,320
 - (B) FILING DATE: May 11, 1999
 - (C) CLASSIFICATION:
- (vii) PRIOR APPLICATION DATA:
 - (A) APPLICATION NUMBER: 09/309,320
 - (B) FILING DATE: May 11, 1999
 - (A) APPLICATION NUMBER: 09/096,571
 - (B) FILING DATE: June 12, 1998
 - (A) APPLICATION NUMBER: 08/756,771
 - (B) FILING DATE: November 26, 1996
- (viii) ATTORNEY/AGENT INFORMATION:
 - (A) NAME: Billings, Lucy J.
 - (B) REGISTRATION NUMBER: 36,749
 - (C) REFERENCE/DOCKET NUMBER: PF-0162 US
- (ix) TELECOMMUNICATION INFORMATION:
 - (A) TELEPHONE: 415-855-0555
 - (B) TELEFAX: 415-845-4166
 - (C) TELEX:

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 222 amino acids
 - (B) TYPE: amino acid
 - (C) STRANDEDNESS: single
 - (D) TOPOLOGY: linear

- (vii) IMMEDIATE SOURCE:
 (A) LIBRARY: Consensus
 (B) CLONE: Consensus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
Met Ala Ala Arg Pro Lys Leu His Tyr Pro Asn Gly Arg Gly Arg Met
1      5      10      15
Glu Ser Val Arg Trp Val Leu Ala Ala Gly Val Glu Phe Asp Glu
20     25     30
Glu Phe Leu Glu Thr Lys Glu Gln Leu Tyr Lys Leu Gln Asp Gly Asn
35     40     45
His Leu Leu Phe Gln Gln Val Pro Met Val Glu Ile Asp Gly Met Lys
50     55     60
Leu Val Gln Thr Arg Ser Ile Leu His Tyr Ile Ala Asp Lys His Asn
65     70     75
Leu Phe Gly Lys Asn Leu Lys Glu Arg Thr Leu Ile Asp Met Tyr Val
85     90     95
Glu Gly Thr Leu Asp Leu Leu Glu Leu Leu Ile Met His Pro Phe Leu
100    105    110
Lys Pro Asp Asp Gln Gln Lys Glu Val Val Asn Met Ala Gln Lys Ala
115    120    125
Ile Ile Arg Tyr Phe Pro Val Phe Glu Lys Ile Leu Arg Gly His Gly
130    135    140
Gln Ser Phe Leu Val Gly Asn Gln Leu Ser Leu Ala Asp Val Ile Leu
145    150    155
Leu Gln Thr Ile Leu Ala Leu Glu Glu Lys Ile Pro Asn Ile Leu Ser
165    170    175
Ala Phe Pro Phe Leu Gln Glu Tyr Thr Val Lys Leu Ser Asn Ile Pro
180    185    190
Thr Ile Lys Arg Phe Leu Glu Pro Gly Ser Lys Lys Lys Pro Pro Pro
195    200    205
Asp Glu Ile Tyr Val Arg Thr Val Tyr Asn Ile Phe Arg Pro
210    215    220
```

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 800 base pairs
 (B) TYPE: nucleic acid
 (C) STRANDEDNESS: single
 (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: Consensus
 (B) CLONE: Consensus

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
ATGGCAGCAA GGCCTCAAGCT CCACTATCCC AACGGAAGAG GCCGGATGGA GTCCGTGAGA 60
TGGSTTTTAG CTCGCGCCGG AGTCGAGTTT GATGAAGAAT TTCTGGAAAC AAAAGAACAG 120
TTGTACAAGT TGCAGGATGG TAACCCAGCTG CTGTTCCAAC ATCTGCCCAT GGTTGAAATT 180
GACGGGATGA AGTTGGTACA GACCCGAAGC ATTCTCCACT ACATAGCAGA CAAGCACAAT 240
CTCTTTGGCA AGAACCTCAA GGAGAGAAAC CTGATTGACA TGTACGTGGA GGGGACACTG 300
GATCTGCTGG AACTGCTTAT CATGCATCCT TTCTTAAAC CAGATGATCA GCAAAAGGAA 360
GTGGTTAACA TGGCCAGAA GCCTATAATT AGATACTTTC CTGTGTTTGA AAAGATTTTA 420
AGGGGTACAG GACAAAGCTT TCTTGTGGT AATCAGCTGA GCCTGTCAGA TGTGATTTTA 480
CTCCAAACCA TTTTAGCTCT AGAAGAGAAA ATTCTTAATA TCCTGTCTGC ATTTCCTTTC 540
CTCCAGGAAT ACACAGTGAA ACTAAGTAAT ATCCCTACAA TTAAGAGATT CTTTGAACCT 600
GGCAGCAAGA AGAAGCCTCC CCCTGATGAA ATTTATGTGA GAACCGTCTA CAACATCTTT 660
AGGCCATAAA ACAACACATC CATGTGTGAG TGACAGTGTG TTCCTAGAGA TGTATTGTC 720
```

TACAGTCATG TCTTAATGGA TCCCAGCTCT GTCATGGTGC TATCTATGTA TTAAGTTGGG
TCCTAAGTTG GGTCCTTTGT

780
800

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 222 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: GenBank
- (B) CLONE: 825605

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```

Met Ala Glu Lys Pro Lys Leu His Tyr Ser Asn Thr Arg Gly Arg Met
 1           5           10           15
Glu Ser Ile Arg Trp Leu Leu Ala Ala Gly Val Glu Phe Glu Glu
 20           25           30
Lys Phe Ile Lys Ser Ala Glu Asp Leu Asp Lys Leu Arg Asn Asp Gly
 35           40           45
Tyr Leu Met Phe Gln Gln Val Pro Met Val Glu Ile Asp Gly Met Lys
 50           55           60
Leu Val Gln Thr Arg Ala Ile Leu Asn Tyr Ile Ala Ser Lys Tyr Asn
 65           70           75           80
Leu Tyr Gly Lys Asp Ile Lys Glu Lys Ala Leu Ile Asp Met Tyr Ile
 85           90           95
Glu Gly Ile Ala Asp Leu Gly Glu Met Ile Leu Leu Leu Pro Phe Thr
100           105           110           115
Gln Pro Glu Glu Gln Asp Ala Lys Leu Ala Leu Ile Gln Glu Lys Thr
120           125
Lys Asn Arg Tyr Phe Pro Ala Phe Glu Lys Val Leu Lys Ser His Gly
130           135           140
Gln Asp Tyr Leu Val Gly Asn Lys Leu Ser Arg Ala Asp Ile His Leu
145           150           155           160
Val Glu Leu Leu Tyr Tyr Val Glu Glu Leu Asp Ser Ser Leu Ile Ser
165           170           175
Ser Phe Pro Leu Leu Lys Ala Leu Lys Thr Arg Ile Ser Asn Leu Pro
180           185           190
Thr Val Lys Lys Phe Leu Gln Pro Gly Ser Pro Arg Lys Pro Pro Met
195           200           205
Asp Glu Lys Ser Leu Glu Glu Ser Arg Lys Ile Phe Arg Phe
210           215           220

```

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 222 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: GenBank
- (B) CLONE: 259141

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Ala Glu Lys Pro Lys Leu His Tyr Phe Asn Ala Arg Gly Arg Met

1	5	10	15
Glu Ser Thr Arg Trp Leu Leu Ala Ala Gly Val Glu Phe Glu Glu	20	25	30
Lys Phe Ile Lys Ser Ala Glu Asp Leu Asp Lys Leu Arg Asn Asp Gly	35	40	45
Tyr Leu Met Phe Gln Gln Val Pro Met Val Glu Ile Asp Gly Met Lys	50	55	60
Leu Val Gln Thr Arg Ala Ile Leu Asn Tyr Ile Ala Ser Lys Tyr Asn	65	70	75
Leu Tyr Gly Lys Asp Ile Lys Glu Arg Ala Leu Ile Asp Met Tyr Ile	85	90	95
Glu Gly Ile Ala Asp Leu Gly Glu Met Ile Leu Leu Leu Pro Val Cys	100	105	110
Pro Pro Glu Lys Asp Ala Lys Leu Ala Leu Ile Lys Glu Lys Ile	115	120	125
Lys Asn Arg Tyr Phe Pro Ala Phe Glu Lys Val Leu Lys Ser His Gly	130	135	140
Gln Asp Tyr Leu Val Gly Asn Lys Leu Ser Arg Ala Asp Ile His Leu	145	150	155
Val Glu Leu Leu Tyr Tyr Val Glu Glu Leu Asp Ser Ser Leu Ile Ser	165	170	175
Ser Phe Pro Leu Leu Lys Ala Leu Lys Thr Arg Ile Ser Asn Leu Pro	180	185	190
Thr Val Lys Lys Phe Leu Gln Pro Gly Ser Pro Arg Lys Pro Pro Met	195	200	205
Asp Glu Lys Ser Leu Glu Glu Ala Arg Lys Ile Phe Arg Phe	210	215	220

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 222 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(vii) IMMEDIATE SOURCE:

- (A) LIBRARY: GenBank
- (B) CLONE: 193710

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met Ala Ala Lys Pro Lys Leu Tyr Tyr Phe Asn Gly Arg Gly Arg Met	1	5	10	15
Glu Ser Ile Arg Trp Leu Leu Ala Ala Gly Val Glu Phe Glu Glu	20	25	30	
Glu Phe Leu Glu Thr Arg Glu Gln Tyr Glu Lys Met Gln Lys Asp Gly	35	40	45	
His Leu Leu Phe Gly Gln Val Pro Leu Val Glu Ile Asp Gly Met Met	50	55	60	
Leu Thr Gln Thr Arg Ala Ile Leu Ser Tyr Leu Ala Ala Lys Tyr Asn	65	70	75	80
Leu Tyr Gly Lys Asp Leu Lys Glu Arg Val Arg Ile Asp Met Tyr Ala	85	90	95	
Asp Gly Thr Gln Asp Leu Met Met Met Ile Ala Val Ala Pro Phe Lys	100	105	110	
Thr Pro Lys Glu Lys Glu Glu Ser Tyr Asp Leu Ile Leu Ser Arg Ala	115	120	125	
Lys Thr Arg Tyr Phe Pro Val Phe Glu Lys Ile Leu Lys Asp His Gly	130	135	140	
Glu Ala Phe Leu Val Gly Asn Gln Leu Ser Trp Ala Asp Ile Gln Leu				

PF-0162-3 DIV

145					150					155				160
Leu	Glu	Ala	Ile	Leu	Met	Val	Glu	Glu	Leu	Ser	Ala	Pro	Val	Leu
				165					170					175
Asp	Phe	Pro	Leu	Leu	Gln	Ala	Phe	Lys	Thr	Arg	Ile	Ser	Asn	Ile
			180					185					190	
Thr	Ile	Lys	Phe	Leu	Gln	Pro	Gly	Ser	Gln	Arg	Lys	Pro	Pro	Pro
		195				200					205			
Asp	Gly	Pro	Tyr	Val	Glu	Val	Val	Arg	Ile	Val	Leu	Lys	Phe	
	210					215					220			